1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

Journalism is a category in which all the projects done are canceled due to some or the other reason – so the probability of any future project being successful is very low whereas on the other hand music has highest probability of getting a project success

There are some sub - categories – in which all the projects have been successful – such as documentary and rock (having huge number of projects as well) – so any future projects getting success in this sub category have high probability of succeeding.

In the months of April and May there are relatively slightly higher chances of succeeding, however it doesn’t make a lot of difference on the probability of getting success. Although we can make a conclusion that as the year progresses the funds available with Kickstarter keep decreasing and hence lesser projects get funded.

1. What are some limitations of this dataset?

We can see that the sample data of 4000 projects is not a good representative set of 300,000 projects since the problem statement mentions 1/3rd of the total 300k projects giving a positive outcome while our analysis shows a success rate of 54% which is clearly a lot more.

Also there is a lot of US data i.e. 3000 (sheet 7) odd projects out of 4000 which again may not represent the 300k projects.

1. What are some other possible tables and/or graphs that we could create?

Sheet 7 and Sheet 8 – have other graphs that we could create for getting more insights regarding the country wise distribution of projects and goals vs sub-category vs state graph.